Unveiling environmental entanglement in strongly dissipative qubits

Soumya Bera,[1] Serge Florens,[1] Harold U. Baranger,[2] Nicolas Roch,[3] Ahsan Nazir,[4] and Alex W. Chin[5]

[1] *Institut Néel, CNRS and UJF, B.P. 166, 25 Avenue des Martyrs, 38042 Grenoble, France*
[2] *Department of Physics, Duke University, Durham, North Carolina 27708, USA*
[3] *Laboratoire Pierre Aigrain, École Normale Supérieure, CNRS (UMR 8551), Université Pierre et Marie Curie, Université Denis Diderot, 24 rue Lhomond, 75231 Paris Cedex 05, France*
[4] *Blackett Laboratory, Imperial College London, London SW7 2AZ, United Kingdom*
[5] *Theory of Condensed Matter Group, University of Cambridge, J J Thomson Avenue, Cambridge, CB3 0HE, United Kingdom*

(Dated: February 1, 2013)

The coupling of a qubit to a macroscopic reservoir plays a fundamental role in understanding the complex transition from the quantum to the classical world. Considering a harmonic environment, we use both intuitive arguments and numerical many-body quantum tomography to study the structure of the complete wavefunction arising in the strong-coupling regime, reached for intense qubit-environment interaction. The resulting strongly-correlated many-body ground state is built from quantum superpositions of adiabatic (polaron-like) and non-adiabatic (antipolaron-like) contributions from the bath of quantum oscillators. The emerging Schrödinger cat environmental wavefunctions can be described quantitatively via simple variational coherent states. In contrast to qubit-environment entanglement, we show that non-classicality and entanglement among the modes in the reservoir are crucial for the stabilization of qubit superpositions in regimes where standard theories predict an effectively classical spin.

The study of dissipative quantum phenomena, namely the interaction of a quantum object (a qubit) with an infinite number of environmental degrees of freedom, lies at the frontier of modern science and technology, with deep implications for fundamental quantum physics[1], quantum computing[2], and even biology[3,4]. While quantum information stored in the qubit subsystem is lost during the coupling with the unobserved degrees of freedom in the reservoir, it is in principle preserved in the entangled many-body state of the global system. The precise nature of this complete wavefunction has received little attention, especially regarding the entanglement generated among the reservoir states. Our purpose here is to unveil a simple emerging structure of the wavefunctions in open quantum systems, using a complementary combination of numerical many-body quantum tomography and a novel analytical variational theory.

An archetype for quantitatively exploring the quantum dissipation problem[5–7] is to start with the simplest quantum object, a two-level system describing a generic quantum bit embodied by spin states $\{|\uparrow\rangle, |\downarrow\rangle\}$, and to couple it to an environment consisting of an infinite collection of quantum oscillators $a_k^\dagger$ (with continuous quantum number k and energy $\hbar\omega_k$). Quantum superposition of the two qubit states is achieved through a splitting Δ acting on the transverse spin component, while dissipation (energy exchange with the bosonic environment) and decoherence are provided by a longitudinal interaction term g_k with each displacement field in the bath. This leads to the Hamiltonian of the celebrated continuum spin-boson model (SBM)[5,6]:

$$H = \frac{\Delta}{2}\sigma_x - \sigma_z \sum_k \frac{g_k}{2}(a_k^\dagger + a_k) + \sum_k \omega_k a_k^\dagger a_k, \qquad (1)$$

where we set $\hbar = 1$, and the sums can be considered as integrals by introducing the spectral function of the environment, $J(\omega) \equiv \sum_k g_k^2 \delta(\omega - \omega_k)$. The generality of the SBM makes it a key model for studying non-equilibrium dynamics, non-Markovian quantum evolution, biological energy transport, and the preparation and control of exotic quantum states in a diverse array of physical and chemical systems[4–8].

The possibility of maintaining robust spin superpositions in the ground and steady states of the SBM has attracted considerable attention, primarily due to its implications for quantum computing[9,10]. Previous numerical approaches have hitherto mainly focused on observables related to the qubit degrees of freedom[11–18], whilst a description of the global system-environment wavefunction has been confined to simpler variational studies[19–23]. This variational theory readily predicts the formation of semiclassical polaron states, which involve the adiabatic response of the environmental modes to the spin tunneling. Strong entanglement between the qubit and the bath is generated in this process. We shall demonstrate here that the many-body ground state of Hamiltonian (1) contains additional non-classical correlations *among* the environmental oscillator modes arising from their *non-adiabatic* response to the spin-flip processes. These new non-classical contributions to the wavefunction are key for the actual stabilization of qubit superpositions relative to the semiclassical picture, and naturally emerge from a variational framework beyond the adiabatic polaron approximation.

In order to enlighten the nature of these emergent non-classical environmental states, we first analyse the SBM by performing the (unitary) polaron transformation $\widetilde{H} = UHU^\dagger$, where $U = \exp\{-\sigma_z \sum_k \frac{g_k}{2\omega_k}(a_k^\dagger - a_k)\}$,

which removes the linear interaction term in Eq. (1). This transforms the Hamiltonian to a basis in which oscillator wavefunctions are displaced according to the z-axis projection of the spin:

$$\widetilde{H} = \frac{\Delta}{2}\sigma^{+}e^{-\sum_k \frac{g_k}{\omega_k}(a_k^\dagger - a_k)} + \text{h.c.} + \sum_k \omega_k a_k^\dagger a_k - E_R, \quad (2)$$

where $E_R = \sum_k g_k^2/(4\omega_k)$ is the reorganisation energy of the bath. For $\Delta = 0$, the ground state of $\widetilde{H}$ is doubly degenerate, and is given by the product of the bosonic vacuum and the spin states, $|\widetilde{\Psi}_{\uparrow,0}\rangle = |\uparrow\rangle \otimes |0\rangle$ and $|\widetilde{\Psi}_{\downarrow,0}\rangle = |\downarrow\rangle \otimes |0\rangle$, in the transformed basis (denoted by tildes). It thus corresponds to polaronic wavefunctions in the original frame, where the positive/negative sign of the displacement is fully correlated to the spin projection (*adiabatic* response): $|\Psi_{\uparrow,g_k/2\omega_k}\rangle = |\uparrow\rangle \otimes |+g_k/2\omega_k\rangle$ and $|\Psi_{\downarrow,-g_k/2\omega_k}\rangle = |\downarrow\rangle \otimes |-g_k/2\omega_k\rangle$. The two-fold degenerate ground state thus takes the form of a product of semiclassical coherent states (displaced oscillators) $|\pm f_k\rangle \equiv e^{\pm\sum_k f_k(a_k^\dagger - a_k)}|0\rangle$, with displacements $f_k = \pm g_k/2\omega_k$ which shift each oscillator to the minimum of its static spin-dependent potential. This potential is evident in Eq. (1) for $\Delta = 0$ and is shown explicitly in Fig. 1**A**. In the presence of spin tunneling ($\Delta \neq 0$), one needs to understand the effect of the operators $K_\pm \equiv \Delta\sigma^{\pm}e^{\mp\sum_k (g_k/\omega_k)(a_k^\dagger - a_k)}$ in Eq. (2) which correlate spin flip processes with simultaneous displacements of *all* oscillator states. As we now show, these correlations ultimately control the ground state qubit superposition.

Polarons, antipolarons, and ground state ansatz. The optimum oscillator displacements result from a competition between the two terms appearing in Hamiltonian (2), namely spin tunneling Δ versus oscillator kinetic energy. Within the transformed frame, consider the coupling induced by the tunneling operator $K_\pm$ between one of the doubly degenerate states, say $|\widetilde{\Psi}_{\downarrow,0}\rangle = |\downarrow\rangle \otimes |0\rangle$, and a spin-flipped state with arbitrary displacement function $\widetilde{f}_k \equiv f_k - g_k/2\omega_k$, $|\widetilde{\Psi}_{\uparrow,\widetilde{f}_k}\rangle = |\uparrow\rangle \otimes |\widetilde{f}_k\rangle$ (f_k is the displacement in the original frame). The matrix element for this is

$$\langle\widetilde{\Psi}_{\uparrow,\widetilde{f}_k}|K_+|\widetilde{\Psi}_{\downarrow,0}\rangle = \Delta e^{-\frac{1}{2}\sum_k(\widetilde{f}_k + g_k/\omega_k)^2}. \quad (3)$$

The elastic displacement energy of oscillator k in $|\widetilde{\Psi}_{\uparrow,\widetilde{f}_k}\rangle$ is $\langle\widetilde{\Psi}_{\uparrow,\widetilde{f}_k}|\omega_k a_k^\dagger a_k|\widetilde{\Psi}_{\uparrow,\widetilde{f}_k}\rangle = \omega_k \widetilde{f}_k^2$. The polaron transformed Hamiltonian reveals the inherent competition between the (elastic) energetic cost of mixing displaced oscillators into the ground state, favouring $\widetilde{f}_k = 0$, and the exponential suppression of the spin kinetic energy, given by the reduced tunneling matrix element in Eq. (3), which rather favours $\widetilde{f}_k = -g_k/\omega_k$. For high-frequency modes, the elastic energy cost dominates and tunneling between spin states is governed by environment states with $\widetilde{f}_k = 0$, gaining only a small (renormalized) tunneling energy $\Delta_R = \Delta e^{-\frac{1}{2}\sum_k (g_k/\omega_k)^2} \ll \Delta$ for strong qubit-environment interaction. In the original frame, the corresponding displacement is $f_k = +g_k/2\omega_k$, which implies that these 'fast' oscillators instantaneously (adiabatically) tunnel with the spin between the minima of their elastic potentials – see Fig. 1. In the opposite limit of low-frequency modes ($\omega_k \ll \Delta_R$), the elastic energy barrier is weak; mixing between spin states is instead governed by the matrix element (3). Returning to the original frame, one gets an energy gain of the *bare* tunneling energy Δ when $f_k = -g_k/2\omega_k$. As shown in Fig. 1**C**, this corresponds to spin tunneling with *non-adiabatic response* of the oscillators, which are displaced in the *opposite* direction from the adiabatic modes. We naturally dub these contributions to the wavefunction *antipolaron states*. At intermediate frequencies, we expect that both polaronic and antipolaronic responses occur, leading to a two-polaron ansatz for the ground state (in the original frame):

$$\begin{aligned}\left|GS^{2\text{pol.}}\right\rangle = &\left|\uparrow\right\rangle \otimes \left[\left|+f_k^{\text{pol.}}\right\rangle + p\left|+f_k^{\text{anti.}}\right\rangle\right] \\ &- \left|\downarrow\right\rangle \otimes \left[\left|-f_k^{\text{pol.}}\right\rangle + p\left|-f_k^{\text{anti.}}\right\rangle\right],\end{aligned} \quad (4)$$

with p the relative weight of the polaron and antipolaron components. Note that this ansatz fully respects the symmetries of the Hamiltonian.

This state reduces to standard (adiabatic) polaron theory when $p = 0$ and $f_k^{\text{pol.}} = g_k/2\omega_k$, and to the variational polaron state of Silbey and Harris (SH)[19,20] when $p = 0$ and the function $f_k^{\text{pol.}}$ is varied to minimise the total ground state energy $E = \langle GS^{2\text{pol.}}|H|GS^{2\text{pol.}}\rangle$. As we shall compare our ansatz (4) to these simpler theories, a brief description of them is given in the Supplementary Information. For $p \neq 0$, the environment wave function for each spin projection is a multi-modal Schrödinger cat state involving a superposition of polaronic and antipolaronic components, leading to considerable mode entanglement. The critical observation is that such superposition of displaced states lowers the energy of the ground state by *stabilising* the spin energy. For the state (4) the spin tunneling energy $E_T = (\Delta/2)\langle GS^{2\text{pol.}}|\sigma_x|GS^{2\text{pol.}}\rangle$ is

$$\begin{aligned}E_T = &-\Delta e^{-2\sum_k (f_k^{\text{pol.}})^2} - p^2\Delta e^{-2\sum_k (f_k^{\text{anti.}})^2} \\ &- 2p\Delta e^{-\frac{1}{2}\sum_k (f_k^{\text{pol.}} + f_k^{\text{anti.}})^2}.\end{aligned} \quad (5)$$

The first two terms reflect an exponentially suppressed renormalized tunneling rate. Indeed, for strong coupling, the displacements $f_k^{\text{pol.}}$ and $f_k^{\text{anti.}}$ are large, and the associated contribution to the spin energy becomes vanishingly small. However, the overlap between the polaron and antipolaron contributions (the third term) will not be suppressed if, as we expect, $f_k^{\text{anti.}} \approx -f_k^{\text{pol.}}$. The development of a small but finite antipolaron weight p thus allows the environment to minimise its displacement energy whilst maintaining significant overlap between the environment-dressed spin states.

Single mode. Before tackling the challenging many-mode situation, we develop intuition about the polaron-

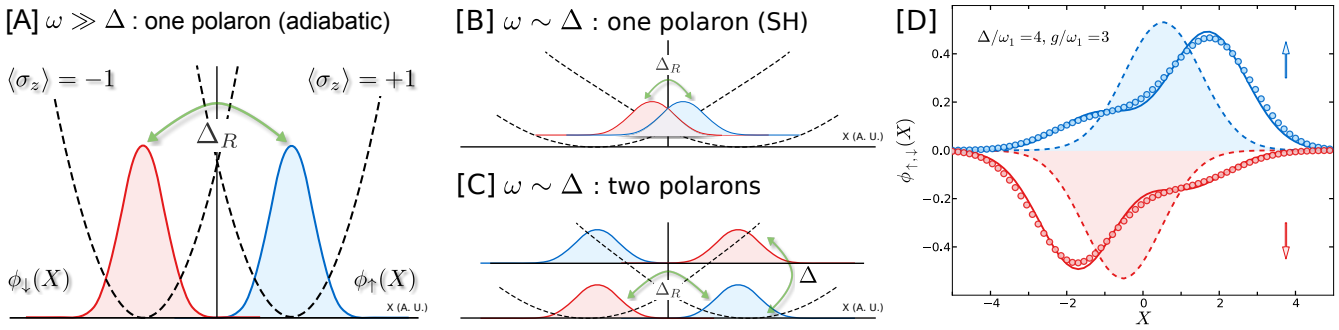

FIG. 1. **Origins of polaron and antipolaron displacements in environment wavefunctions.** In plots **A-C**, black dashed lines are the spin-dependent potential energies of a single harmonic oscillator in the absence of spin tunneling [see Eq. (1)], while blue (red) curves are the gaussian wavefunctions (in real space x) of the oscillator on the $\langle\sigma_x\rangle = 1$ (-1) potential surfaces. **A. Polarons**. For a high frequency mode ($\omega \gg \Delta$), transitions to other oscillator states on the potential surfaces are suppressed by the steep curvature of the potentials; oscillator displacement adiabatically tunnels with the spin between minima of the potentials, suppressing the tunneling amplitude by the reduced overlap of the displaced oscillator wave functions to a value Δ_R. **B. Non-adiabatic response in Silbey-Harris variational polaron theory**. Low frequency modes ($\omega \ll \Delta$) have shallow potentials, leading to well-separated minima. Poor wave function overlap prevents tunneling of the spin between minima, destroying spin superposition. Variationally-determined displacements adjust to smaller values, sacrificing their displacement energy to maintain the spin-tunneling energy through better overlap. **C. Antipolaron response of non-adiabatic oscillators**. For modes with $\omega \sim \Delta$, spin flips that do not change the position of the oscillators (and thus have unsuppressed amplitude Δ) may become low enough in energy to compete with the overlap-suppressed inter-minima tunneling. The oscillator wavefunctions correlated with spin are now superpositions of displaced coherent states with opposite signs. **D. Ground state wavefunction components of a single oscillator**. Spin up (blue) and spin down (red) components are shown for the exact ground state (circles), our variational polaron-antipolaron state (solid lines), and the Silbey-Harris ansatz (dashed lines) [$\Delta/\omega_1 = 4$, $g/\omega_1 = 3$]. The exact result shows distinct antipolaron features which are well captured by the variational polaron-antipolaron state. The Silbey-Harris ansatz shows reduced displacements and thus poor agreement with the exact result.

antipolaron ansatz in the simplest case of a single environmental mode with energy ω_1 and coupling g_1. This case is easily diagonalised numerically (see also Ref. 24 for an exact solution); note that a similar ansatz for the single-mode Rabi model (without reference to polaron theory) has been previously explored numerically[25,26]. In Figure 1**D** we compare the spatial wavefunctions of the oscillator correlated with each spin state with those obtained from the ansatz Eq. (4) following a numerical optimization of p, $f_1^{\text{pol.}}$, and $f_1^{\text{anti.}}$ to minimise the ground state energy. Choosing oscillator parameters where we expect non-adiabatic response, namely $\omega_1 < \Delta$, we find that both wavefunctions clearly show a superposition of polaron and antipolaron contributions, with much larger displacements compared to the prediction of the SH theory (single polaron case $p = 0$). The agreement of the diagonalised and the two-polaron ansatz ground state wavefunctions is extremely good, as well as the energies and spin observables, even for a coupling strength as large as $g = 3\omega_1$ (see also Supplementary Information). As motivated above, the emergence of an antipolaron component in the environment enhances the overlap of the tunneling states. The single polaron SH state fails in this regard (see Figures 1**B** and 1**D**, and Supplementary Information) as it finds itself frustrated between minimizing the elastic energy and maintaining good overlap between the opposite spin states: the resulting displacements are thus totally wrong.

Two-mode antipolaron entanglement. Having confirmed the emergence of non-adiabatic antipolaron contributions in the case of a single mode, we now consider the case of a two-mode SBM and, in particular, test our proposal Eq. (4) that the two-mode wave function dressing a given spin state will be entangled.

Fig. 2 shows the spin-up component of the two-mode wavefunction as a function of the two independent spatial coordinates of the modes (x_1 and x_2) for two modes taken at different frequencies $\omega_2 = 2\omega_1$. The ground state wavefunctions were determined by exact numerical diagonalisation. We see the clear development of an antipolaron component to the wavefunction (Fig. 2**B**) for low-energy non-adiabatic modes, in contrast to the situation of high-energy adiabatic modes (Fig. 2**A**). However, we see that only two peaks appear in the wavefunction – those along the diagonal line $x_1 = x_2$ – indicating unambiguously that this two-mode wavefunction takes the inter-mode entangled form $|f_1^{\text{pol.}}\rangle \otimes |f_2^{\text{pol.}}\rangle + p|f_1^{\text{anti.}}\rangle \otimes |f_2^{\text{anti.}}\rangle$. This can be contrasted with a hypothetical polaron-antipolaron product state $\{|f_1^{\text{pol.}}\rangle + p|f_1^{\text{anti.}}\rangle\} \otimes \{|f_2^{\text{pol.}}\rangle + p|f_2^{\text{anti.}}\rangle\}$ which would rather display four peaks, as shown in Fig. 2**C**. The implications of this inter-mode entanglement for the entropy of the reservoir modes is given in Supplementary Information. Again, one can check that the variational energy of the

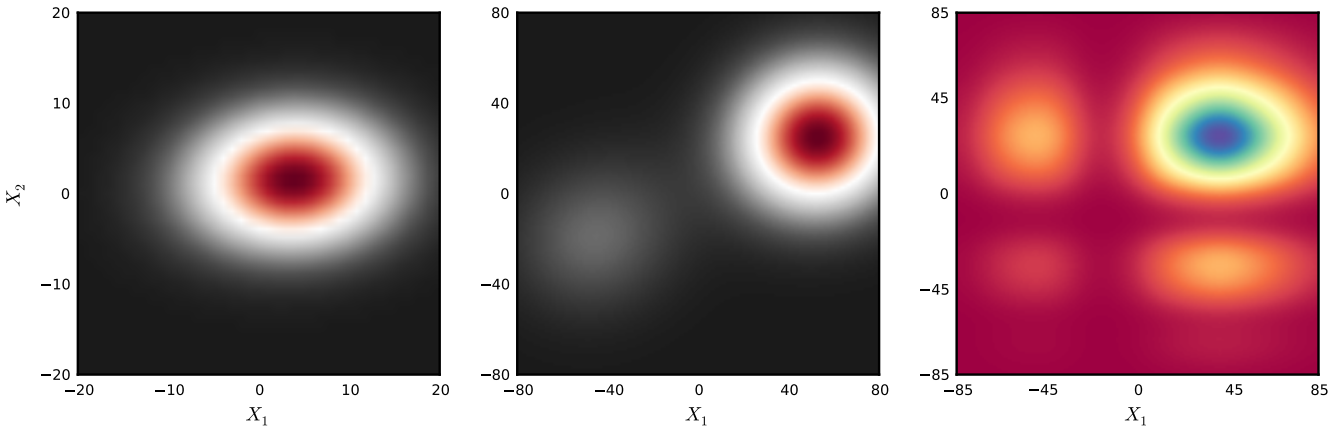

FIG. 2. **Two mode wavefunctions**. **A-B**. Contour plots in real space of the spin-up projected joint oscillator wavefunctions of two modes obtained from exact diagonalisation. **A. Polarons.** For high frequency modes ($\omega_2 = 2\omega_1 = 0.04 > \Delta = 0.01$), the wavefunction is a single, displaced gaussian, in qualitative agreement with Silbey-Harris theory. **B. Entangled antipolarons.** Low frequency modes ($\omega_2 = 2\omega_1 = 0.004 < \Delta = 0.01$) show the development of an antipolaron component, visible in the ($X_1 < 0, X_2 < 0$) quadrant, in addition to the Silbey-Harris state. **C. Product state.** Hypothetical wavefunction obtained from a product state of polaron-antipolaron superpositions for each mode, showing symmetric off-diagonal peaks. These features are absent in **B**, indicating that the exact joint wavefunction is not a product state but, in contrast, is entangled as described in the text.

two-mode ground state is remarkably close to the exact energy.

Multi-mode spin-boson model. We now turn to the more challenging many-mode situation, tackling the continuum spin-boson model (1). A direct diagonalisation of the Hamiltonian is now hopeless; however, recent computational progress has opened the way to calculating ground state averages of arbitrary operators, for instance using the bosonic Numerical Renormalization Group (NRG)[27], which we will use to test the generalized polaron state (4). A key feature in the NRG method is the use of a logarithmic discretization of the energy spectrum of the bath, which ensures the stability and convergence of an iterative diagonalization of the impurity model[11]. In order to directly compare with the variational results, we use the same discretization in defining the polaronic ansatz Eq. (4), incorporating the changing measure in ω_k into the definitions of the f_k.

We focus here on the standard case of ohmic dissipation[5,6], although our following results should apply similarly to other types of spectral density. The continuous bath of bosonic excitations assumes then a linear spectrum in frequency, $J(\omega) = 2\alpha\omega\theta(\omega_c - \omega)$, up to a high energy cutoff ω_c and with dimensionless dissipation strength α. Weakly damped Rabi oscillations of the qubit for $\alpha \ll 1$ are known to completely fade away in the strong dissipation regime $\alpha \gtrsim 0.4$, where the qubit becomes strongly entangled with its environment. The bare qubit frequency Δ is heavily renormalized in this regime to the smaller value $\Delta_R = \Delta(\Delta e/\omega_c)^{\alpha/(1-\alpha)}$, for $\Delta/\omega_c \ll 1$, which can thus be driven to zero for the critical dissipation strength $\alpha_c \simeq 1$, indicating a quantum critical point.

As a first step towards understanding the many-mode situation, we consider the variational solution obtained from the two-polaron ansatz (4) (the variational equations are given in the Supplementary Information). This leads to the polaronic and antipolaronic displacements shown in Fig. 3**A**, which exemplify the physical picture introduced above (see especially Fig. 1): polaron and antipolaron states show equal and opposite displacements at low energies (typically for $\omega_k \ll \omega_c$), but merge together to produce a fully polaronic state at high energy, where the environment responds adiabatically to the spin. The variational theory is thus able, without additional physical input, to generate the correct crossover from non-adiabatic to adiabatic behavior of the antipolaron component with increasing energy.

The presence of the antipolaron component has a large impact on the ground state spin average: Fig. 3**B** compares the result of the one- and two-polaron variational states to that computed with NRG (numerically exact result, used as a benchmark). In the one-polaron (SH) limit ($p = 0$), one finds readily $-\langle\sigma_x\rangle = \Delta_R/\Delta = (\Delta e/\omega_c)^{\alpha/(1-\alpha)}$, which incorrectly vanishes at the critical dissipation strength $\alpha_c = 1$[19,22]. On the other hand, the emergence of antipolaron correlations at low energy, namely $f_k^{\text{anti.}} \simeq -f_k^{\text{pol.}}$ for $\omega_k \ll \omega_c$, helps in maintaining a finite value for $\langle\sigma_x\rangle$, due to the perfect cancellation of the displacements within the exponential in the last term of Eq. (5). This success of the antipolaron ansatz (4) is illustrated in Fig. 3**B**.

Our objective now is to demonstrate the peculiar inter-mode entanglement properties of the antipolaron ansatz (4). While one cannot plot the complete many-body wavefunction in the case of many environmental

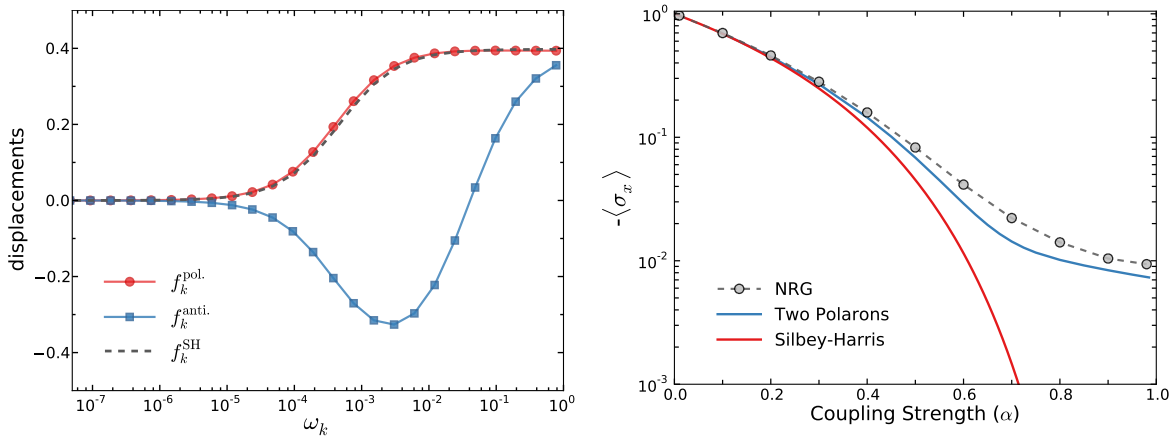

FIG. 3. **Displacements and spin average in the many-mode case. A.** Displacements determined variationally from the two-polaron ansatz Eq. (4), showing the emergence of an antipolaron component for low energies, with equal and opposite displacement to the polaron state. The antipolaron state merges smoothly onto the polaron state at high energy as the adiabaticity of the oscillators with respect to tunneling of the spin is recovered (the NRG logarithmic discretisation of the bath spectrum is used here, namely frequency points are evenly spaced on a logarithmic scale; Note that a point at higher energy is associated to a larger energy window of the continuum spectrum, leading to the saturation of $f_k^{\text{pol.}}$ for high frequencies, instead of the fall off obtained for a linear energy mesh). [Parameters: $\alpha = 0.5$ and $\Delta = 0.01$.] **B.** Ground state averaged spin amplitude $-\langle\sigma_x\rangle$ as a function of dissipation strength α computed with the NRG (circles) for $\Delta/\omega_c = 0.01$, and compared to the one-polaron (red line) and two-polaron (blue line) predictions. A clear breakdown of the one-polaron Silbey-Harris ansatz occurs at strong dissipation, while the two-polaron trial state accounts for the correct behavior up to the quantum critical point ($\alpha_c = 1$), due to preserved tunneling amplitude via the antipolaron component of the wavefunction.

modes, a useful strategy to assess the validity of the trial state (4) lies in recent interest in quantum tomography[28–30], wherein the reduced density matrix in a smaller projected Hilbert space is fully characterized. For the problem at hand, we trace out all modes except the qubit degree of freedom together with an *arbitrary* bath mode with given quantum number k; this defines a spin and k-mode excluded environment denoted "env/spin+k". The reduced ground state density matrix in the joint qubit and k-mode subspace reads

$$\rho_{\text{spin}+k} = \text{Tr}_{\text{env/spin}+k}|GS\rangle\langle GS|. \tag{6}$$

We focus here on the off-diagonal part (with respect to the qubit axis of quantization) of the Wigner distribution associated to this density matrix as a function of the classical displacement X. We expect on physical grounds that this component will be most sensitive to the antipolaronic correlations. Its standard definition is[1]

$$W^{(k)}_{\sigma^+}(X) = \int \frac{\mathrm{d}^2\lambda}{\pi^2}\, e^{X(\lambda-\bar{\lambda})}\text{Tr}_{\text{spin}+k}\left[e^{\lambda a_k^\dagger - \bar{\lambda}a_k}\sigma^+\rho_{\text{spin}+k}\right]; \tag{7}$$

see Methods for the NRG implementation and Supplementary Information for discussion of the spin-diagonal part of the Wigner distribution, which emphasizes instead the polaronic part of the total wavefunction. From the two-polaron trial state (4) and equation (7), it is straightforward to find the form of the Wigner function in the regime of strong dissipation ($\alpha > 0.5$):

$$W^{(k)}_{\sigma^+}(X) \approx \frac{p}{\pi} e^{-\frac{1}{2}\sum_{q\neq k}(f_q^{\text{pol.}}+f_q^{\text{anti.}})^2} \times \left[e^{-2\left(X-\frac{f_k^{\text{pol.}}-f_k^{\text{anti.}}}{2}\right)^2} + e^{-2\left(X+\frac{f_k^{\text{pol.}}-f_k^{\text{anti.}}}{2}\right)^2}\right]. \tag{8}$$

For high energy modes $\omega_k \sim \omega_c$, $W^{(k)}_{\sigma^+}(X)$ should show a single peak centered around $X = 0$, as both polarons adiabatically follow the spin tunneling so that $f_k^{\text{anti.}}$ becomes close to the polaron displacement $f_k^{\text{pol.}}$. For modes of lower energy, antipolaron displacements emerge, and the peak separates into two lobes with displacements $\pm[f_k^{\text{pol.}} - f^{\text{anti.}}] \simeq \pm 2 f_k^{\text{pol.}}$. These simple predictions of the two-polaron variational state are clearly seen in the numerical NRG result in Fig. 4, strongly supporting the existence of non-adiabatic oscillator states in the environment.

We finally wish to assess more directly the entanglement among the environmental states that is suggested by the antipolaron ansatz Eq. (4). For this purpose, we consider the entanglement entropy $S_{\text{spin}+k}$ of the joint spin and k-mode subsystem with respect to the other modes of the bath:

$$S_{\text{spin}+k} = -\text{Tr}_{\text{spin}+k}\left[\rho_{\text{spin}+k}\log\rho_{\text{spin}+k}\right], \tag{9}$$

with the reduced density matrix defined in Eq. (6). We also introduce the spin entanglement entropy $S_{\text{spin}} = -\text{Tr}_{\text{spin}}\left[\rho_{\text{spin}}\log\rho_{\text{spin}}\right]$ with $\rho_{\text{spin}} = \text{Tr}_k[\rho_{\text{spin}+k}]$. The difference between these two quantities can be computed

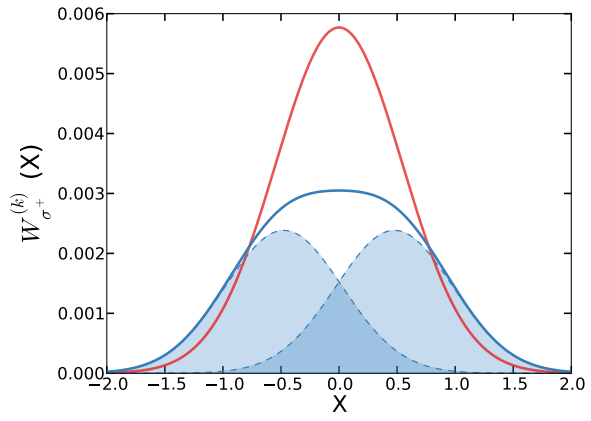

FIG. 4. **Quantum tomography in the many-mode case.** Transverse Wigner distribution defined in Eq. (7), as obtained from NRG, for two different modes, one at high energy $\omega_k \gg \Delta_R$ (top red curve) and the other at intermediate energy $\omega_k \gtrsim \Delta_R$ (bottom blue curve). A decomposition of the intermediate energy case into two shifted Gaussians is performed (thin blue lines) according to Eq. (8), showing the emergence of antipolaron correlations. [Parameters $\alpha = 0.8$ and $\Delta/\omega_c = 0.01$.]

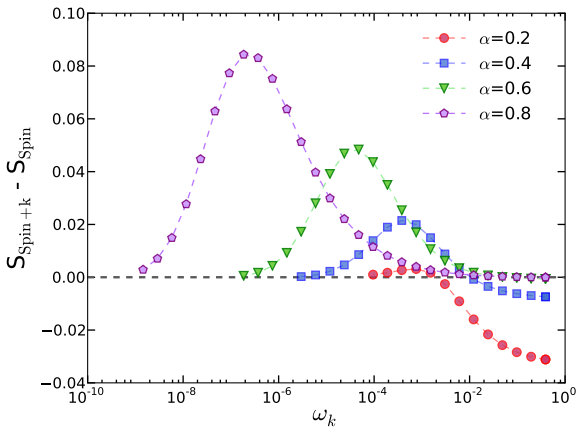

FIG. 5. **Joint entanglement entropy in the many-mode case.** The joint entanglement entropy of the qubit and a given ω_k mode, defined by Eq. (9), is calculated with NRG for $\Delta/\omega_c = 0.01$. The entropy of the qubit alone is subtracted. Negative correlations for $\alpha < 0.5$ are related to the Silbey-Harris one-polaron state, while the strong positive peak can only be accounted for by entanglement among the modes of the bath. This excess entropy is, then, a sensitive measure of the subtle non-classical correlations among the bath modes that are generated by the coupling to the qubit.

from the NRG (see Methods) and is plotted as a function of mode frequency in Fig. 5. For small dissipation, $\alpha < 0.5$, this entropy difference is mostly negative, as expected from the correlations built into the Silbey-Harris state, which consist only of non-entangled environmental states within each spin-projected component of the wavefunction (see Supplementary Information). In contrast, at strong dissipation, this entropy difference becomes positive and shows a strikingly large enhancement near the scale Δ_R. The excess entanglement entropy, above that of the spin alone, comes from entanglement within the bath of oscillators. This is a sensitive signature, then, that the spin projected wavefunction is not simply a product of oscillator states as in the SH ansatz but rather involves substantial entanglement. The nature of this entanglement in the simpler two-mode case is explored further in the Supplementary Information. Note especially the large energy window where the entropy peak develops: the excess entanglement spreads from low to high frequency modes due to the massive entangling power of the spin tunneling operator K_+ discussed above in the polaron basis $\widetilde{H}$ of Eq. (1). The existence of inter-mode bosonic correlations on a wide energy range makes also an interesting connection to the underlying (although hidden in the spin-boson model) fermionic Kondo physics[5,31].

In conclusion, we have shown how antipolaron contributions emerge in the ground state wavefunction of the spin-boson model, causing non-classical Schrödinger cat-like environmental states. The approach here can be used as a general framework to expand and rationalize many-body wavefunctions in strongly interacting open quantum systems. Experimentally, proposals to realize the strongly dissipative spin boson model using a superconducting qubit coupled to arrays of Josephson junctions have been made very recently[31–33]. The recent progress in quantum tomography of superconducting qubits[29,30] raises thus the challenge to measure in such setups the massive entanglement of environment oscillators that was unveiled here. The present work offers several directions for future research, especially the generalization to time-dependent phenomena such as the study of quantum quenches and spin dynamics at strong dissipation, where standard (weak-coupling) Bloch-Redfield theory[7] is known to fail.

Methods

The numerical solution of the few-mode spin-boson Hamiltonian (1) relies on standard diagonalisation procedures. In the case of a continuous bath of oscillators, a different strategy is used. First, logarithmic shell blocking of the bosonic modes onto energy intervals $[\Lambda^{-n-1}\omega_c, \Lambda^{-n}\omega_c]$ with $\Lambda = 2$ is performed:

$$a_n^\dagger = \int_{\Lambda^{-n-1}\omega_c}^{\Lambda^{-n}\omega_c} \mathrm{d}k \; a_k^\dagger \; . \tag{10}$$

The resulting discrete Hamiltonian, which spans from arbitrarily small energy up to the high energy cutoff ω_c, is then iteratively diagonalised according to the Numerical Renormalization Group (NRG) algorithm[11,27]. The novel part of the simulations performed for this work lies in the computation of the Wigner distribution reduced to the joint spin and single k-mode subspace. In order to

implement Eqs. (6-7), we first define arbitrary moments of the chosen oscillator k of frequency $\omega_k = \omega_c \Lambda^{-n}$:

$$A^{(k)}_{\sigma_i;m,m'} = \big\langle GS|\sigma_i[a_n^\dagger]^m[a_n]^{m'}|GS\big\rangle \tag{11}$$

with $i = 0, x, y, z$ labelling the Pauli matrices related to the spin projection (we take $\sigma_0 \equiv 1$) and m, m' positive integers. Such ground state observables are readily computed within the NRG algorithm (for typically $0 \leq m, m' < 10$). One can then expand Eq. (7) in a power series in λ and $\bar{\lambda}$, yielding

$$W^{(k)}_{\sigma_i}(X) = \frac{2}{\pi} \sum_{m,m'=0}^{+\infty} A^{(k)}_{\sigma_i;m,m'} \frac{(-1)^{m+m'}}{m!m'!} \frac{\partial^{m+m'}}{\partial X^{m+m'}} e^{-2X^2}. \tag{12}$$

The Wigner distribution is now solely expressed in terms of the NRG-computable moments $A^{(k)}_{\sigma_i;m,m'}$.

A similar strategy is used for the computation of the entanglement entropy (9) from the reduced density matrix $\rho_{\mathrm{spin}+k}$, which acts within the subspace spanned by the qubit and a single bosonic mode k. We start by defining the joint spin and Fock projection operator $O^{(k)}_{\sigma_i;m,m'} = \sigma_i|m_k\big\rangle\big\langle m'_k|$, so that matrix elements of the ground state density matrix simply read

$$\rho^{(k)}_{\sigma_i,m,m'} = \big\langle GS|O^{(k)}_{\sigma_i;m,m'}|GS\big\rangle\ . \tag{13}$$

This quantity is a ground state average, hence readily computable by letting the operator $O^{(k)}_{\sigma_i;m,m'}$ evolve along the complete NRG flow. The eigenvalues of the matrix $\rho^{(k)}_{\sigma_i,m,m'}$ allow one, finally, to obtain the desired entanglement entropy.

A last new piece of our work is the multipolaron generalization of the previous single-polaron trial state[19,20]; this is key for capturing easily the emergent non-adiabatic physics at strong dissipation. The variational method is straightforwardly implemented in the few-mode case by minimizing the average Hamiltonian (1) while using the double-polaron ansatz (4). In the many mode case, despite having two sets of unknown functions, $f_k^{\mathrm{pol.}}$ and $f_k^{\mathrm{anti.}}$, labeled by the continuous momentum k, one can show that their form as a function of k is uniquely fixed from the variational principle, leaving a finite set of effective parameters to be determined (see Supplementary Information). One finds that the displacement associated with the first polaron follows qualitatively the standard behavior $f_k^{\mathrm{pol.}} = 0.5 g_k/(\omega_k + \Delta_R)$ known from SH theory[19,20], with some quantitative deviations due to the feedback of the antipolaronic state $f_k^{\mathrm{anti.}}$. The latter takes the approximate form $f_k^{\mathrm{anti.}} \simeq f_k^{\mathrm{pol.}} \cdot \frac{\omega_k - \Omega}{\omega_k + \Omega}$ with a new energy scale Ω that controls the crossover from non-adiabatic to adiabatic behavior as a function of mode energy (see Supplementary Information for the complete expression). The antipolaronic (non-adiabatic) character at low energy of the second contribution in the trial state (4) is thus automatically guaranteed by the variational principle.

[1] Raimond, J. M. & Haroche, S. *Understanding the Quantum* (Oxford Graduate Series, 2006).

[2] Nielsen, A. M. & Chuang, I. L. *Quantum Computation and Quantum Information* (Cambridge University Press, New York, 2007).

[3] Lambert, N., Chen, Y.-N., Cheng, Y.-C., Li, C.-J., Chen, G.-Y. & Nori, F. Quantum biology. *Nature Physics* **9**, 10 (2013).

[4] Scholes, G., Fleming, G., Olaya-Castro, A. & van Grondelle, R. *Nature Chemistry* **3**, 763 (2011).

[5] Leggett, A. J., Chakravarty, S., Dorsey, A. T., Fisher, M. P. A., Garg, A. & Zwerger W. Dynamics of the dissipative two-state system. *Rev. Mod. Phys.* **59**, 1 (1987).

[6] Weiss, U. *Quantum Dissipative Systems* (World Scientific, 1993).

[7] Breuer, H.-P. & Petruccione, F. *The Theory of Open Quantum Systems,* (Oxford University Press, 2010).

[8] Nitzan, A. *Chemical Dynamics in Condensed Phases: Relaxation, Transfer and Reactions in Condensed Molecular Systems* (Oxford University Press, 2006).

[9] Jennings, D., Dragan, A., Barrett, S. D., Bartlett, S. D. & Rudolph, T. Quantum computation via measurements on the low-temperature state of a many-body system. *Phys. Rev. A* **80**, 032328 (2009).

[10] Raussendorf, R. & Briegel, H. J. A one-way quantum computer. *Phys. Rev. Lett.* **86**, 5188 (2001).

[11] Bulla, R., Costi, T. A. & Pruschke, T. Numerical renormalization group method for quantum impurity systems. *Rev. Mod. Phys.* **80**, 395 (2008).

[12] Makri, N. Numerical path integral techniques for long time dynamics of quantum dissipative systems. *J. Math. Phys.* **36**, 2430 (1995).

[13] Wang, H. & Thoss, M. From coherent motion to localization: dynamics of the spin-boson model at zero temperature. *New J. Phys.* **10**, 115005 (2008).

[14] Nalbach, P. & Thorwart, M. Ultraslow quantum dynamics in a sub-ohmic heat bath. *Phys. Rev. B* **81**, 054308 (2010).

[15] Winter, A., Rieger, H., Vojta, M. & Bulla, R. Quantum phase transition in the sub-ohmic spin-boson model: Quantum Monte-Carlo study with a continuous imaginary time cluster algorithm. *Phys. Rev. Lett.* **102**, 030601 (2009).

[16] Alvermann, A. & Fehske, H. Sparse polynomial space approach to dissipative quantum systems: Application to the sub-ohmic spin-boson model. *Phys. Rev. Lett.* **102**, 150601 (2009).

[17] Prior, J., Chin, A. W., Huelga, S. F. & Plenio, M. B. Efficient simulation of strong system-environment interactions. *Phys. Rev. Lett.* **105**, 050404 (2010).

[18] Florens, S., Freyn, A., Venturelli, D. & Narayanan, R. Dissipative spin dynamics near a quantum critical point: Numerical renormalization group and Majorana diagrammatics. *Phys. Rev. B* **84**, 155110 (2011).

[19] Silbey, R. & Harris, R. Variational calculation of the dynamics of a two level system interacting with a bath. *J. Chem. Phys.* **80**, 2615 (1984).
[20] Harris, R. A. & Silbey, R. Variational calculation of the tunneling system interacting with a heat bath. II. Dynamics of an asymmetric tunneling system. *J. Chem. Phys.* **83**, 1069 (1985).
[21] Chin, A. W., Prior, J., Huelga, S. F. & Plenio, M. B. Generalized polaron ansatz for the ground state of the sub-ohmic spin-boson model: An analytic theory of the localization transition. *Phys. Rev. Lett.* **107**, 160601 (2011).
[22] Nazir, A., McCutcheon, D. P. S. & Chin, A. W. Ground state and dynamics of the biased dissipative two-state system: Beyond variational polaron theory. *Phys. Rev. B* **85**, 224301 (2012).
[23] Agarwal, K., Martin, I., Lukin, M. D. & Demler, E. Polaronic model of Two Level Systems in amorphous solids. *Preprint, arXiv:1212.3299*.
[24] Braak, D. Integrability of the Rabi model. *Phys. Rev. Lett.* **107**, 100401 (2011).
[25] Hwang, M.-Y. & Choi, M.-S. Variational study of a two-level system coupled to a harmonic oscillator in an ultrastrong-coupling regime. *Phys. Rev. A.* **82**, 025802 (2010).
[26] Stolze, J. & Müller, L. Quality of variational ground states for a two-state system coupled to phonons. *Phys. Rev. B.* **42**, 6704 (1990).
[27] Bulla, R., Tong, N.-H. & Vojta, M. Numerical renormalization group for bosonic systems and application to the sub-ohmic spin-boson model. *Phys. Rev. Lett.* **91**, 170601 (2003).
[28] Lvovsky, A. I. & Raymer, M. G. Continuous-variable optical quantum-state tomography. *Rev. Mod. Phys.* **81**, 299 (2009).
[29] Hofheinz, M., Wang, H., Ansmann, M., Bialczak, R. C., Lucero, E., Neeley, M., O'Connell, A. D., Sank, D., Wenner, J., Martinis, J. M. & Cleland, A. N. Synthesizing arbitrary quantum states in a superconducting resonator. *Nature* **459**, 546 (2009).
[30] Eichler, C., Bozyigit, D., Lang, C., Steffen, L., Fink, J. & Wallraff, A. Experimental state tomography of itinerant single microwave photons. *Phys. Rev. Lett.* **106**, 220503 (2011).
[31] Le Hur, K. Kondo resonance of a microwave photon. *Phys. Rev. B* **85**, 140506 (2012).
[32] Ballester, D., Romero, G., García-Ripoll, J. J., Deppe, F. & Solano, E. Quantum simulation of the ultrastrong-coupling dynamics in circuit quantum electrodynamics. *Phys. Rev. X* **2**, 021007 (2012).
[33] Goldstein, M., Devoret, M. H., Houzet, M. & Glazman, L. I. Inelastic microwave photon scattering off a quantum impurity in a Josephson-junction array. *Phys. Rev. Lett.* **110**, 017002 (2013).

Acknowledgements

S.B., S.F., and H.U.B. thank the Fondation Nanosciences de Grenoble for funding under RTRA contract CORTRANO. A.N. thanks Imperial College for support. A.W.C. acknowledges support from the Winton Programme for the Physics of Sustainability. The work at Duke was supported by US DOE, Division of Materials Sciences and Engineering, under Grant No. desc0005237.

Supplementary information for "Unveiling environmental entanglement in strongly dissipative qubits"

We present here additional technical details and extra results, supporting the multi-polaronic description of the many-body ground state of the spin-boson model at strong coupling. We first consider the general two-polaron variational formalism for an arbitrary number of modes. The ground state energy and wavefunctions are then investigated for a wide range of parameter in the single-mode (Rabi) model, highlighting the emergence of antipolaron correlations and the possible breakdown of the single-polaron Silbey-Harris ansatz. The two-mode Rabi model is afterwards considered, with the emphasis on entropic issues, which provide interesting signatures of environmental entanglement. The need for additional antipolaronic contributions in the wavefunction is also discussed. Finally, the continuous spin-boson model is further explored, with detailed derivations of the Wigner functions pertaining to the reduced qubit and single mode Hilbert space, as well as extra comparisons between Numerical Renormalization Group simulations and the variational technique.

I. GENERAL TWO-POLARON VARIATIONAL FORMALISM

A. Energetics

We consider the unbiased spin-boson model[1,2], as defined by the Hamiltonian (1) of the main text:

$$H = \frac{\Delta}{2}\sigma_x + \sum_k \omega_k a_k^\dagger a_k - \sigma_z \sum_k \frac{g_k}{2}(a_k^\dagger + a_k), \tag{1}$$

with tunneling energy Δ, a set of oscillator frequencies ω_k, and system-oscillator coupling strengths g_k (assumed real). Here, $\sigma_z = |\uparrow\rangle\langle\uparrow| - |\downarrow\rangle\langle\downarrow|$, with spin basis states $|\downarrow\rangle$ and $|\uparrow\rangle$, and $a_k^\dagger$ (a_k) is the oscillator creation (annihilation) operator for mode k. Hamiltonian (1) spans the cases of few discrete modes up to a continuum of bosonic fields, in which case the discrete k-sum ought to be replaced by an integral over energy.

Our two-polaron variational ground state ansatz takes the form

$$\left|GS^{2\text{pol.}}\right\rangle = |\uparrow\rangle\left[p_1\left|+f_k^{\text{pol.}}\right\rangle + p_2\left|+f_k^{\text{anti.}}\right\rangle\right] - |\downarrow\rangle\left[p_1\left|-f_k^{\text{pol.}}\right\rangle + p_2\left|-f_k^{\text{anti.}}\right\rangle\right], \tag{2}$$

where the bosonic part of the wavefunction involves coherent states of the form

$$|\pm f_k\rangle = e^{\pm\sum_k f_k(a_k^\dagger - a_k)}|0\rangle, \tag{3}$$

defined as products of displaced states, where $|0\rangle$ represents all oscillators being in the vacuum state. The presence of a $\mathbb{Z}_2$ symmetry, namely ($|\uparrow\rangle \to |\downarrow\rangle$, $|\downarrow\rangle \to |\uparrow\rangle$, $a_k \to -a_k$), and the need for minimizing the spin tunneling energy enforces the chosen relative sign between the up and down components of the ground state wavefunction in Eq. (2). Both functions $f_k^{\text{pol.}}$ and $f_k^{\text{anti.}}$ are taken as free parameters, and will be varied to minimise the total ground state energy $E = \left\langle GS^{2\text{pol.}}\right| H \left|GS^{2\text{pol.}}\right\rangle$. In contrast to the usual Silbey-Harris state (for which $p_2 = 0$)[3–5], this more flexible ansatz allows for the possibility of a superposition of variationally determined displaced oscillator states associated with each spin projection.

Normalisation of $\left|GS^{2\text{pol.}}\right\rangle$ implies the condition

$$2p_1^2 + 2p_2^2 + 4p_1p_2e^{-\frac{1}{2}\sum_k (f_k^{\text{pol.}} - f_k^{\text{anti.}})^2} = 1, \tag{4}$$

while the variational ground state energy is given by

$$\begin{aligned} E = \left\langle GS^{2\text{pol.}}\right| H \left|GS^{2\text{pol.}}\right\rangle = &-\Delta\left(p_1^2 e^{-2\sum_k (f_k^{\text{pol.}})^2} + p_2^2 e^{-2\sum_k (f_k^{\text{anti.}})^2} + 2p_1p_2e^{-\frac{1}{2}\sum_k (f_k^{\text{pol.}} + f_k^{\text{anti.}})^2}\right) \\ &+2\sum_k \omega_k\left(p_1^2 (f_k^{\text{pol.}})^2 + p_2^2 (f_k^{\text{anti.}})^2 + 2p_1p_2 f_k^{\text{pol.}} f_k^{\text{anti.}} e^{-\frac{1}{2}\sum_k (f_k^{\text{pol.}} - f_k^{\text{anti.}})^2}\right) \\ &-2\sum_k g_k\left(p_1^2 f_k^{\text{pol.}} + p_2^2 f_k^{\text{anti.}} + p_1p_2(f_k^{\text{pol.}} + f_k^{\text{anti.}})e^{-\frac{1}{2}\sum_k (f_k^{\text{pol.}} - f_k^{\text{anti.}})^2}\right). \end{aligned} \tag{5}$$

In the limit that $p_2 \to 0$ (and so $p_1 \to 1/\sqrt{2}$) we recover the Silbey-Harris variational ground state energy,

$$E_{\text{SH}} = -\frac{\Delta}{2}e^{-2\sum_k (f_k^{\text{pol.}})^2} + \sum_k \omega_k (f_k^{\text{pol.}})^2 - \sum_k g_k f_k^{\text{pol.}}, \tag{6}$$

while further setting $\alpha_k = g_k/(2\omega_k)$ in Eq. (6) gives the (non-variationally optimal) bare polaron ground state energy

$$E_{\text{POL}} = -\frac{\Delta}{2}e^{-\frac{1}{2}\sum_k g_k^2/\omega_k^2} - \sum_k \frac{g_k^2}{4\omega_k}. \tag{7}$$

Going back to the two-polaron variational state of Eq. (2), we find that the ground state coherence is given by

$$\langle\sigma_x\rangle = -2\left(p_1^2 e^{-2\sum_k (f_k^{\text{pol.}})^2} + p_2^2 e^{-2\sum_k (f_k^{\text{anti.}})^2} + 2p_1p_2e^{-\frac{1}{2}\sum_k (f_k^{\text{pol.}} + f_k^{\text{anti.}})^2}\right), \tag{8}$$

while the magnetisation $\langle\sigma_z\rangle = 0$ by symmetry in absence of magnetic field along σ_z (unless one enters the polarized phase at $\alpha > 1$ in the ohmic spin-boson model).

B. Variational displacements

The two sets of displacements $f_k^{\text{pol.}}$ and $f_k^{\text{anti.}}$ are variationally determined from the total energy E of Eq. (5) according to $\partial E/\partial f_k^{\text{pol.}} = 0$ and $\partial E/\partial f_k^{\text{anti.}} = 0$, which gives the closed form:

$$f_k^{\text{pol.}} = \frac{g_k}{2}\frac{A_1(p_2^2\omega_k + \Delta_2) - A_2[p_1p_2\langle f_k^{\text{pol.}}|f_k^{\text{anti.}}\rangle + \Delta_{12}]}{(p_1^2\omega_k + \Delta_1)(p_2^2\omega_k + \Delta_2) - [p_1p_2\langle f_k^{\text{pol.}}|f_k^{\text{anti.}}\rangle + \Delta_{12}]^2} \tag{9}$$

$$f_k^{\text{anti.}} = \frac{g_k}{2}\frac{A_2(p_1^2\omega_k + \Delta_1) - A_1[p_1p_2\langle f_k^{\text{pol.}}|f_k^{\text{anti.}}\rangle + \Delta_{12}]}{(p_1^2\omega_k + \Delta_1)(p_2^2\omega_k + \Delta_2) - [p_1p_2\langle f_k^{\text{pol.}}|f_k^{\text{anti.}}\rangle + \Delta_{12}]^2}, \tag{10}$$

which is valid for an arbitrary number of oscillator modes. Hence the generic k-dependence of the displacement is fully constrained by the variational principle, which leaves a finite set of effective parameters to be determined self-consistently according to:

$$\Delta_1 = \Delta p_1^2 \langle - f_k^{\text{pol.}} | f_k^{\text{pol.}} \rangle + \frac{\Delta}{2} p_1 p_2 \langle - f_k^{\text{pol.}} | f_k^{\text{anti.}} \rangle + p_1 p_2 (-\tilde{\omega} + \tilde{g}) \langle f_k^{\text{pol.}} | f_k^{\text{anti.}} \rangle \tag{11}$$

$$\Delta_2 = \Delta p_2^2 \langle - f_k^{\text{anti.}} | f_k^{\text{anti.}} \rangle + \frac{\Delta}{2} p_1 p_2 \langle - f_k^{\text{pol.}} | f_k^{\text{anti.}} \rangle + p_1 p_2 (-\tilde{\omega} + \tilde{g}) \langle f_k^{\text{pol.}} | f_k^{\text{anti.}} \rangle \tag{12}$$

$$\Delta_{12} = \frac{\Delta}{2} p_1 p_2 \langle - f_k^{\text{pol.}} | f_k^{\text{anti.}} \rangle + p_1 p_2 (\tilde{\omega} - \tilde{g}) \langle f_k^{\text{pol.}} | f_k^{\text{anti.}} \rangle \tag{13}$$

$$A_1 = p_1^2 + 2 p_1 p_2 \langle f_k^{\text{pol.}} | f_k^{\text{anti.}} \rangle \tag{14}$$

$$A_2 = p_2^2 + 2 p_1 p_2 \langle f_k^{\text{pol.}} | f_k^{\text{anti.}} \rangle \tag{15}$$

$$\tilde{\omega} = \sum_k \omega_k f_k^{\text{pol.}} f_k^{\text{anti.}} \tag{16}$$

$$\tilde{g} = \sum_k g_k [f_k^{\text{pol.}} + f_k^{\text{anti.}}]. \tag{17}$$

The one-polaron Silbey-Harris displacement $f_k^{\text{SH}} = 0.5 g_k / \left[\omega_k + \Delta \langle - f_k^{\text{pol.}} | f_k^{\text{pol.}} \rangle\right]$ is trivially recovered from Eq. (9) by letting $p_2 = 0$. One can also check that $f_k^{\text{anti.}} \simeq f_k^{\text{pol.}}$ for $k \to \infty$, while $f_k^{\text{anti.}} \simeq -f_k^{\text{pol.}}$ for $k \to 0$ in the limit of strong dissipation. Thus the antipolaron displacement satisfies the expected adiabatic/non-adiabatic crossover as a function of energy ω_k, and this physical picture is naturally incorporated in the variational theory.

II. SINGLE-MODE RABI MODEL: CHECKING THE WAVEFUNCTION

It is illustrative to consider the simplified case of a single-mode within the environment, namely the Rabi model (see Ref. 6 and references therein). In this situation, the model Hamiltonian may be diagonalised straightforwardly (numerically) and so the regimes of validity of our polaron-antipolaron ansatz, as well as of the Silbey-Harris and non-optimal bare polaron states, may be assessed. The Hamiltonian now becomes

$$H_1 = \frac{\Delta}{2} \sigma_x + \omega_1 a_1^\dagger a_1 - \sigma_z \frac{g}{2} (a_1^\dagger + a_1), \tag{18}$$

with ground-state ansatz

$$\left| GS_1^{\text{2pol.}} \right\rangle = |\uparrow\rangle \left[p_1 \left| +f_1^{\text{pol.}} \right\rangle + p_2 \left| +f_1^{\text{anti.}} \right\rangle \right] - |\downarrow\rangle \left[p_1 \left| -f_1^{\text{pol.}} \right\rangle + p_2 \left| -f_1^{\text{anti.}} \right\rangle \right], \tag{19}$$

where $|\pm f_1\rangle = e^{\pm f_1 (a_1^\dagger - a_1)} |0\rangle$. To optimise the state, we minimise the variational ground state energy $E = \left\langle GS_1^{\text{2pol.}} \right| H_1 \left| GS_1^{\text{2pol.}} \right\rangle$ numerically, subject to the normalisation constraint $2p_1^2 + 2p_2^2 + 4 p_1 p_2 e^{-\frac{1}{2}(f_1^{\text{pol.}} - f_1^{\text{anti.}})^2} = 1$. The Silbey-Harris state is again obtained by letting $p_2 \to 0$, such that

$$\left| GS_1^{\text{SH}} \right\rangle = \frac{1}{\sqrt{2}} \left[|\uparrow\rangle \left| +f_1^{\text{SH}} \right\rangle - |\downarrow\rangle \left| -f_1^{\text{SH}} \right\rangle \right], \tag{20}$$

and there is only a single displaced oscillator associated with each spin. Minimisation of $E^{\text{SH}} = \left\langle GS_1^{\text{SH}} \right| H_1 \left| GS_1^{\text{SH}} \right\rangle$, leads to a self-consistent equation for the optimised displacement, $f_1^{\text{SH}} = g / [2(\Delta e^{-2(f_1^{SH})^2} + \omega_1)]$. The non-optimal bare polaron state has the same form as Eq. (20), but with the displacement fixed at $f_1 = g/(2\omega_1)$.

In Fig. 1 we plot the dimensionless ground state energy E/ω_1 determined from our polaron-antipolaron variational ansatz as a function of the dimensionless spin-oscillator coupling strength g/ω_1, and compare with the results from an exact numerical diagonalisation of the model, and from Silbey-Harris and polaron theories (Eqs. (6) and (7) in the single mode case, respectively). In this figure, $\Delta/\omega_1 \geq 1$ for all plots, and so we would expect standard polaron theory to break down in this regime, since the full oscillator displacement is no longer appropriate. From the dashed curves, this can indeed be seen to be the case, and polaron theory may even predict the incorrect trend as g/ω_1 increases. Silbey-Harris theory fixes this problem to a certain extent (at least at small g/ω_1, see dashed-dotted curves), though again runs into problems as the coupling strength increases, deviating from the numerically-exact results, and even more worryingly predicting discontinuous behaviour in the ground-state energy at certain values of g/ω_1. Our polaron-antipolaron variational ansatz, however, predicts ground-state energies in almost *perfect agreement* with the numerical

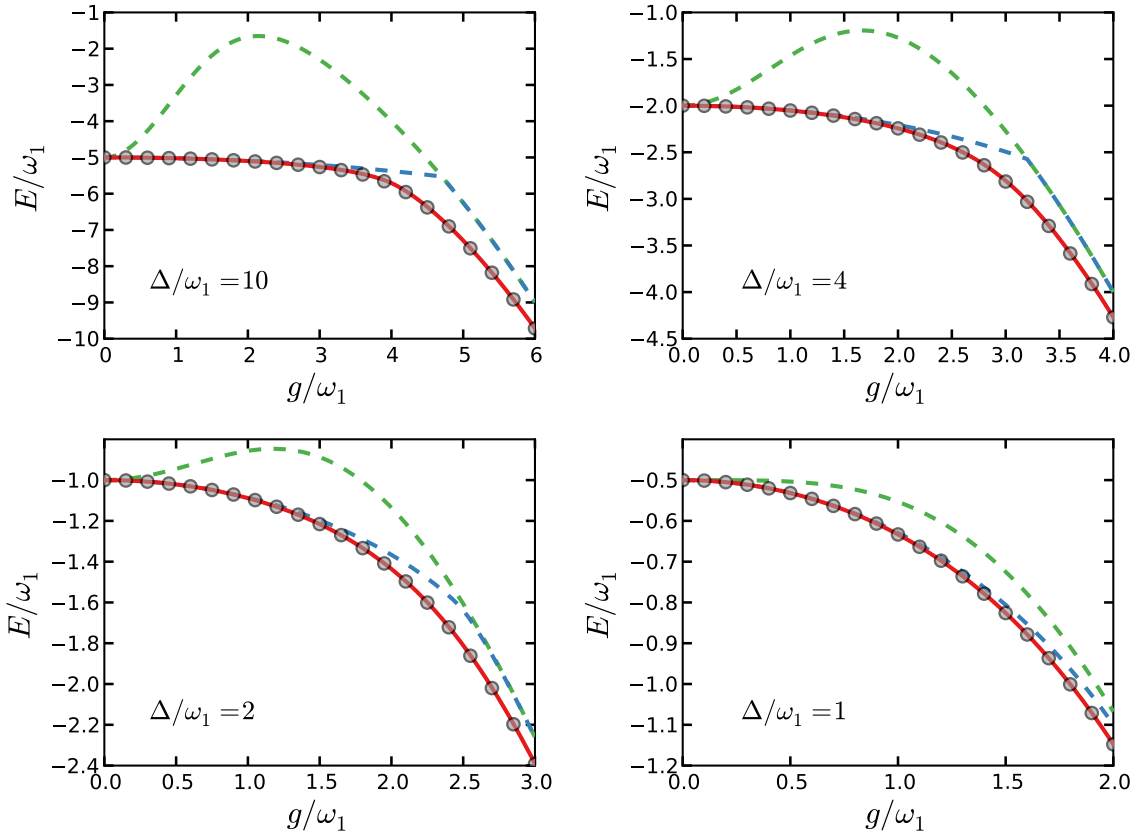

FIG. 1. **Ground state energy of the Rabi model.** The single-mode spin boson model is considered as a function of spin-oscillator coupling strength, and the ground state energy calculated from our variational two-polaron ansatz (red solid curves), the Silbey-Harris one-polaron variational state (blue dashed curves), and from the bare polaron ground state (green dashed curves). Exact energies calculated by numerically diagonalising the Hamiltonian (for a basis of 200 states) are shown as grey dots. The values of Δ/ω_1 used are shown on each plot. For all plots, $\Delta/\omega_1 \geq 1$, which is the regime in which we expect our variational state to outperform the Silbey-Harris or bare polaron treatments.

results for all possible coupling strengths. Furthermore, the discontinuous behaviour seen in Silbey-Harris theory is removed in this more flexible variational state.

The failure of the single-polaron theories can be interpreted by analysing, in position space, the oscillator states associated with the spin projetions $|\downarrow\rangle$ and $|\uparrow\rangle$ in the model ground state ($\phi^0_\downarrow(x)$ and $\phi^0_\uparrow(x)$, respectively), where we take the ground state to have the form $|GS_1\rangle = \left|\phi^0_\uparrow\right\rangle|\uparrow\rangle + \left|\phi^0_\downarrow\right\rangle|\downarrow\rangle$, i.e. we absorb any normalisation factors and minus signs into the oscillator states, $\left|\phi^0_\downarrow\right\rangle$ is thus negative using this convention. Shown in Fig. 2 are thus $\phi^0_\downarrow(x)$ (red) and $\phi^0_\uparrow(x)$ (blue) as a function of position, calculated from our polaron-antipolaron ground state (solid curves), from the Silbey-Harris ground state (dashed-dotted curves), and from a numerical diagonalisation of the full Hamiltonian (points). We take $\Delta/\omega_1 = 4$ here as a representative example. For $g/\omega_1 \sim 1$, the displacement from $x = 0$ is found to be fairly small, $|f| < g/(2\omega_1)$, which is why the full polaron approach fails, see Fig. 1). The correct displacements can be captured by the Silbey-Harris theory and as well by the more flexible two-polaron ansatz presented in Eq. (19). However, as the coupling strength increases further, we can see that the double displacement nature of the oscillator states starts to become extremely important. For example, at $g/\omega_1 = 3$ we observe that the Silbey-Harris state is completely unable to reproduce the correct oscillator wavefunctions, due to the restriction to a single displacement associated with each spin state. In fact, for these parameters, the displacements obtained by the Silbey-Harris approach are much too small, and reproduce none of displacements seen in our polaron-antipolaron ansatz, which itself matches the numerical solution very well. Finally, as the coupling strength is increased further, the Silbey-Harris displacements eventually "jump" to those of the full polaron transformation, which then captures the dominant displacements in the exact states quite well (see the $g/\omega_1 = 4$ plot), but of course completely misses

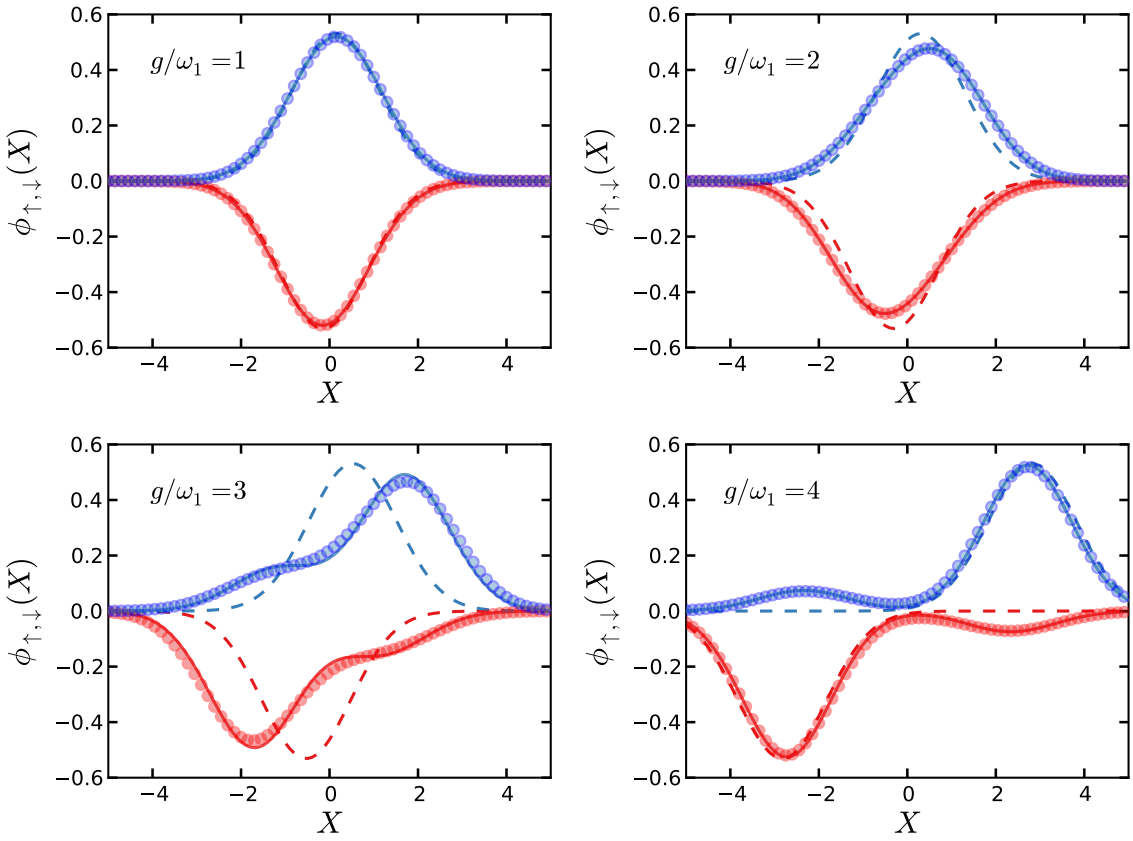

FIG. 2. **Ground state wavefunctions of the Rabi model.** Oscillator states $\phi^0_\downarrow(x)$ (red) and $\phi^0_\uparrow(x)$ (blue) plotted as a function of position for $\Delta/\omega_1 = 4$. Shown are the predictions of our variational ground state (solid curves), of the Silbey-Harris ground state (dashed-dotted curves), and exact calculations from a numerical diagonalisation of the full Hamiltonian (points).

the smaller displacements in the opposite direction, which are still captured extremely well by our ansatz. Hence, the obtained ground state energy is still lower in our ansatz (and in the numerical diagonalisation) than from the Silbey-Harris state. The theories will eventually converge at larger coupling, however, when associating a single (polaron) displacement with each spin state finally becomes a good description.

III. TWO-MODE RABI MODEL: ENTANGLEMENT ENTROPY

In order to understand the non-monotoic behaviour of the joint spin-mode entropy $S_{\text{Spin}+k}$ results presented for the continuum environment in Fig. 5 of the main text, we consider in this appendix the entanglement *between* modes in the environment for the simpler case of a two-mode environment. The Hamiltonian for this case, the same used to generate the two-mode results in the main text, is given by

$$H = \frac{\Delta}{2}\sigma_x - \frac{\sigma_z}{2}\sum_{i=1,2} g_i(a_i + a_i^\dagger) + \sum_{i=1,2}\omega_i a_i^\dagger a_i. \tag{21}$$

For this two-mode environment the ground state can be found by exact diagonalisation (ED) techniques, allowing a comparison to be made with the predictions of our anti-polaron ansatz. Figure 3**A** shows results for $S_{\text{Spin}+1}$ (we trace over mode 2) as a function of ω_1 for ground state obtained by ED and three variational ansätze: Silbey-Harris (one-polaron), two-polaron, and three-polaron (to be discussed below) trial states. In all cases, a fixed relative detuning of the modes and ratio of coupling to frequency was kept, with $\omega_2 = 1.05\omega_1$, $g_1 = g_2 = 2.5\omega_1$. For large frequencies $\omega_1 \gg \Delta_R$, where we expect adiabatic polaron theory to describe the state, the product state (unentangled) form of the wave functions of oscillators 1 and 2 in the spin-projected states is expected to lead $S_{\text{Spin}+1}$ to be controlled only by the polaronic correlations between mode 2 and the spin. This is determined by the renormalisation scale Δ_R, which is

a function of only the ratio g_2/ω_2. As this is held fixed in these simulations, we expect $S_{\mathrm{Spin}+1}$ to approach a constant value at high frequency. This behaviour is indeed observed on all curves in Fig. 3**A**, with the relative strong coupling ($g_2/\omega_2 = 2.5$) leading to an almost fully-mixed spin state. According to Silbey-Harris theory, which has the same wave function structure as the adiabatic polaron theory but with different displacements, oscillators with frequencies well below Δ_R should have suppressed displacements. Consequently, the renormalisation of the spin tunneling by slow modes should be continuously suppressed, leading to reduction of $S_{\mathrm{Spin}+1}$ as $\omega_1 \to 0$. This behaviour is precisely what is observed for the Silbey-Harris results in Fig. 3**A**, with $S_{\mathrm{Spin}+1}$ decreasing monotonically with decreasing frequency (though with a sharp suppression below the spin-tunneling frequency scale).

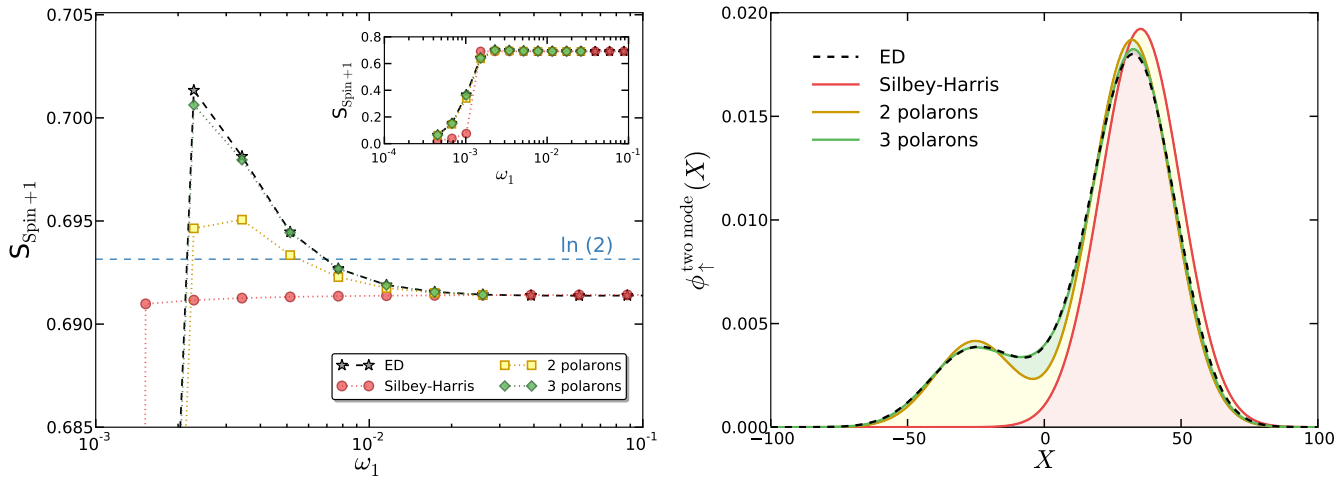

FIG. 3. **Joint spin-mode entropies and oscillator wavefunctions for the two-mode spin-boson model**. **A**. Joint spin-mode entropy $S_{\mathrm{Spin}+1}$ for a two-mode environment. For all these data $\omega_2 = 1.05\omega_1$, $g_1 = g_2 = 2.5\omega_1$ and $\Delta = 0.01$, sweeping the frequencies ω_1 of the first mode. Plot shows results for (top to bottom) exact diagonalisation (grey stars), three-polaron ansatz (green diamonds), two-polaron ansatz (yellow squares) and one-polaron Silbey-Harris theory (red dots). The horizontal line indicates the maximum entropy of a fully-mixed spin state (ln(2)). The main panel is a close-up on the entropy peak occuring near the resonance frequency, as discussed in the text, while the inset shows the whole entropy and frequency range. **B**. Spin up-projected two-mode oscillator wave functions along the diagonal coordinate $x_1 = x_2$ for the same parameters as the left panel, computed for $\omega_1 = 0.0023$ (namely at the peak position of the exact diagonalisation entropy curve in **A**). Results are shown for ground states obtained by exact diagonalisation (dashed black line), one-polaron Silbey-Harris ansatz (solid red line), two-polaron ansatz (solid yellow line) and three-polaron ansatz (solid green line).

However, as in the multimode case, we see that the numerically exact ED ground state shows an entropy peak (in excess of the entropy of a fully-mixed spin state) at frequencies close to the scale Δ_R. As was shown in the two-mode wave function plots in the main text, the spin-projected oscillator wavefunctions in the antipolaron regime are entangled by their non-adiabatic responses, and these inter-mode correlations lead to the entropy peak seen in Fig. 3**A** (due to *both* spin and mode 1 states becoming mixed when mode 2 is traced over). As in the single mode case, this occurs in the regime where antipolarons form, $g_i \approx \omega_i \approx \Delta_R$. The absence of the antipolaron component in the Silbey-Harris theory means that this feature cannot be described by this state.

However, a variationally-optimised two-polaron ground state ansatz as in Eq. (19) captures this peak structure in the frequency dependence of $S_{\mathrm{Spin}+1}$. To understand qualitatively how antipolaron components in the two-polaron ansatz generate inter-mode entanglement and, thus, the increased spin-mode entropy, consider the (spin-up) spin-projected part of the ground state. This gives a contribution to the total density matrix of

$$\rho_{\uparrow\uparrow,1,2} = |\uparrow\rangle\langle\uparrow| \left(p_1^2 |f_k^{\mathrm{pol.}}\rangle\langle f_k^{\mathrm{pol.}}| + p_2^2 |f_k^{\mathrm{anti.}}\rangle\langle f_k^{\mathrm{anti.}}| + p_1 p_2 |f_k^{\mathrm{pol.}}\rangle\langle f_k^{\mathrm{anti.}}| + p_1 p_2 |f_k^{\mathrm{anti.}}\rangle\langle f_k^{\mathrm{pol.}}| \right), \tag{22}$$

where $|f_k^x\rangle = e^{\sum_{k=1,2} f_k^x (a_k^\dagger - a_k)}|0\rangle$. Tracing over mode 2, the reduced state $\rho_{\uparrow\uparrow,1}$ is

$$\rho_{\uparrow\uparrow,1} = |\uparrow\rangle\langle\uparrow| \left(p_1^2 |f_1^{\mathrm{pol.}}\rangle\langle f_1^{\mathrm{pol.}}| + p_2^2 |f_1^{\mathrm{anti.}}\rangle\langle f_1^{\mathrm{anti.}}| + p_1 p_2 \Phi_{22} |f_1^{\mathrm{pol.}}\rangle\langle f_1^{\mathrm{anti.}}| + p_1 p_2 \Phi_{22} |f_1^{\mathrm{anti.}}\rangle\langle f_1^{\mathrm{pol.}}| \right), \tag{23}$$

where $\Phi_{22} = \mathrm{Tr}_2 \left[|f_2^{\mathrm{pol.}}\rangle\langle f_2^{\mathrm{anti.}}| \right] = e^{-(f_2^{\mathrm{pol.}} - f_2^{\mathrm{anti.}})^2/2}$. As $f_2^{\mathrm{pol.}}$ and $f_2^{\mathrm{anti.}}$ have opposite signs when mode 2 is in the

antipolaron regime, the overlap integral Φ_{22} suppresses the purity of the spin-projected state of mode 1. In the extreme case where $\Phi_{22} \approx 0$, the mixed reduced state is

$$\rho_{\uparrow\uparrow,1} \approx |\uparrow\rangle\langle\uparrow| \left(p_1^2 |f_1^{\text{pol.}}\rangle\langle f_1^{\text{pol.}}| + p_2^2 |f_1^{\text{anti.}}\rangle\langle f_1^{\text{anti.}}| \right). \tag{24}$$

This expression provides important intuition as to frequency dependence of inter-mode entanglement. The state given in Eq. (24) will have high entropy if the antipolaron weight p_2/p_1 in the wave function is significant and the states $|f_1^{\text{anti.}}\rangle$ and $|f_1^{\text{pol.}}\rangle$ have weak overlap (i.e. are close to orthogonal) . We therefore expect that the entropy peak will appear in the strong antipolaron regime (where p_2/p_1 and the displacements $f_1^{\text{pol.}} \approx -f_1^{\text{anti.}}$ are sizeable). However, this also requires that Φ_{22} is also ≈ 0, requiring that mode 2 is also in the strong anti-polaron regime. Our theory thus establishes the microscopic link between the appearance of *intra-mode* non-classicality (cat-states/antipolaron) and *many-body* entanglement between such modes. For the results of Fig. 3**A**, the 5% detuning between modes leads to both modes showing antipolaron features at similar frequencies, leading to the large entropy peak. In the multimode case, we therefore expect to find antipolaron components in the environmental wavefunctions everywhere inside the region of positive $S_{\text{Spin}+k} - S_{\text{Spin}}$, allowing the frequency range and position of the non-classical (cat-like) environmental features - which could be experimentally probed through the environmental response function - to be inferred. We have also checked that when the detuning between modes is made larger, such that the modes do not both develop antipolarons for the same parameters, the entropy peak becomes much smaller (not shown).

Finally, we note that while the two-polaron ansatz captures the essential physics of the entropy peak, the preservation of $\langle\sigma_x\rangle$ at strong coupling, and also provides an excellent description of the shapes of the entangled wave functions, the agreement with the ED entropy is not perfect, with the two-polaron ansatz slightly underestimating the peak entropy, see Fig. 3**A**. To investigate this remaining small discrepancy, we have also implemented a variational three-polaron ansatz of the form:

$$\begin{aligned} \left|GS^{3\text{pol.}}\right\rangle = &\left|\uparrow\right\rangle \otimes \left[p_1 \left|+f_k^{\text{pol.}}\right\rangle + p_2 \left|+f_k^{\text{anti.}}\right\rangle + p_3 \left|+f_k^{3}\right\rangle \right] \\ &- \left|\downarrow\right\rangle \otimes \left[p_1 \left|-f_k^{\text{pol.}}\right\rangle + p_2 \left|-f_k^{\text{anti.}}\right\rangle + p_3 \left|-f_k^{3}\right\rangle \right]. \end{aligned} \tag{25}$$

Figure 3**B** shows the wave functions of the spin up-projected states of the oscillators along the diagonal coordinate $x_1 = x_2$ for the ground states obtained by ED and the two and three-polaron ansätze. Comparing the ED and two-polaron wavefunctions, we see that the two-polaron wavefunctions captures the displacements and weights of the polaron and antipolaron very well, but slightly underestimates the amplitude of the wavefunction around the origin. For simplicity, the three-polaron solution was determined by fixing $f_k^3 = 0$ and treating all other parameters variationally. The result, shown in Fig. 3**B** has almost perfect overlap with the ED results and gives an improved prediction for the entropy peak in Fig. 3**A**. This result suggests that it may be fruitful to consider a multipolaron expansion of the state in the many-mode cases, particularly if one is interested in reproducing sensitive measures of the many-body state structure (such as the joint entropy or other tomographic objects) rather than the simple spin observables (which are already well-approximated by the two-polaron results).

IV. MULTI-MODE SPIN-BOSON MODEL: WIGNER DISTRIBUTIONS

We discuss here various Wigner distributions associated to the reduced density matrix living in the subspace spanned by the qubit and one *given* oscillator mode with frequency ω_k. The qubit degrees of freedom can be used for filtering out the polaron and antipolaron contributions within the wavefunction, thanks to appropriate insertions of Pauli matrices in the standard definition of the Wigner function[7]. For instance, we can project onto the $|\uparrow\rangle$ component only, by considering:

$$W^{(k)}_{\frac{1+\sigma^z}{2}}(X) = \int \frac{\mathrm{d}^2\lambda}{\pi^2}\, e^{X(\bar{\lambda}-\lambda)} \left\langle GS \right| e^{\lambda a_k^\dagger - \bar{\lambda} a_k} \frac{1+\sigma^z}{2} \left| GS \right\rangle. \tag{26}$$

This Wigner distribution can be addressed from our two-polaron variational state (2). In the case $\Delta \ll \omega_c$, we find $p_2 \propto \Delta/\omega_c \ll p_1$, and the wavefunction is normalized by taking simply $p_1 \simeq 1/\sqrt{2}$. We thus need to compute:

$$W^{(k)}_{\frac{1+\sigma^z}{2}}(X) = \frac{1}{2} \int \frac{\mathrm{d}^2\lambda}{\pi^2}\, e^{X(\bar{\lambda}-\lambda)} \left[\left\langle f_q^{\text{pol.}} \right| e^{\lambda a_k^\dagger - \bar{\lambda} a_k} \left| f_q^{\text{pol.}} \right\rangle + 2p_2 \left\langle f_q^{\text{pol.}} \right| e^{\lambda a_k^\dagger - \bar{\lambda} a_k} \left| f_q^{\text{anti.}} \right\rangle + p_2^2 \left\langle f_q^{\text{anti.}} \right| e^{\lambda a_k^\dagger - \bar{\lambda} a_k} \left| f_q^{\text{anti.}} \right\rangle \right]. \tag{27}$$

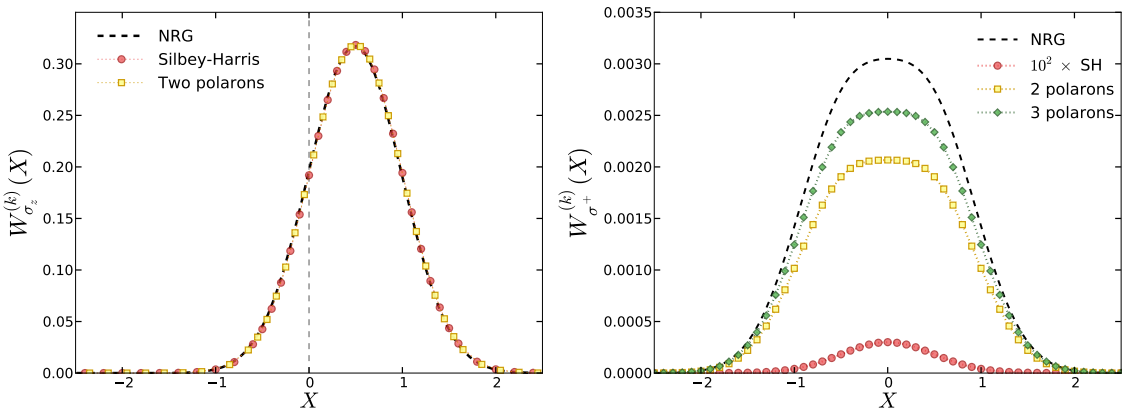

FIG. 4. **Wigner distributions for the many-mode spin-boson model. A.** Diagonal Wigner distribution defined in Eq. (26) computed by the NRG (dashed line) and compared to the one and two-polaron results found from Eq. (31). Parameters are $\alpha = 0.8$ and $\Delta/\omega_c = 0.01$. The one-polaron Silbey-Harris state is enough to account very well for the polaron content of the exact wavefunction. **B.** Off-diagonal Wigner distribution defined in Eq. (33) computed by the NRG (dashed line) and compared to the one and two-polaron results found from Eq. (34). The general magnitude and non-Gaussian form of the Wigner distribution is only accounted for by the two-polaron state, with a complete failure of the Silbey-Harris wavefunction (note the 100× magnification used to reveal its tiny contribution to the Wigner distribution). The remaining quantitative deviations in the two-polaron ansatz are due to the presence of additional antipolaronic contributions in the total wavefunction, as can be inferred from the computation done within a three-polaron trial state.

Using usual coherent state algebra, we readily obtain the required overlaps:

$$\left\langle f_q^{\text{pol.}}\middle|e^{\lambda a_k^\dagger - \bar{\lambda} a_k}\middle|f_q^{\text{pol.}}\right\rangle = e^{(\lambda-\bar{\lambda})f_k^{\text{pol.}}} e^{-\lambda\bar{\lambda}/2} \tag{28}$$

$$\left\langle f_q^{\text{pol.}}\middle|e^{\lambda a_k^\dagger - \bar{\lambda} a_k}\middle|f_q^{\text{anti.}}\right\rangle = e^{-\frac{1}{2}\sum_q (f_q^{\text{pol.}} - f_g^{\text{anti.}})^2} e^{\lambda f_k^{\text{pol.}} - \bar{\lambda} f_k^{\text{anti.}}} e^{-\lambda\bar{\lambda}/2} \tag{29}$$

$$\left\langle f_q^{\text{anti.}}\middle|e^{\lambda a_k^\dagger - \bar{\lambda} a_k}\middle|f_q^{\text{anti.}}\right\rangle = e^{(\lambda-\bar{\lambda})f_k^{\text{anti.}}} e^{-\lambda\bar{\lambda}/2} \tag{30}$$

and performing the Gaussian integral in Eq. (27) yields

$$W^{(k)}_{\frac{1+\sigma^z}{2}}(X) = \frac{1}{\pi} e^{-2(X-f_k^{\text{pol.}})^2} + \frac{2p_2}{\pi} e^{-\frac{1}{2}\sum_{q\neq k}(f_q^{\text{pol.}} - f_q^{\text{anti.}})^2} e^{-2\left(X - \frac{f_k^{\text{pol.}} + f_k^{\text{anti.}}}{2}\right)^2} + \frac{p_2^2}{\pi} e^{-2(X-f_k^{\text{anti.}})^2}. \tag{31}$$

Note that the second term in Eq. (31) provides only a small correction to the first contribution, a feature which follows from (i) $p_2 \ll 1$ and (ii) the overlap appearing in this second contribution can be approximated as $e^{-\frac{1}{2}\sum_{q\neq k}(f_q^{\text{pol.}} - f_q^{\text{anti.}})^2} \simeq \left\langle f_q^{\text{pol.}}\middle|f_q^{\text{anti.}}\right\rangle \simeq e^{-2\sum_q (f_q^{\text{pol.}})^2} \simeq \Delta_R/\Delta \ll 1$ (for $\alpha \gtrsim 0.5$) because the antipolaron is equal and opposite to the polaron at low energy. Similarly, the third term in Eq. (31), which would peak at the antipolaron displacement, is of order $p_2^2 \ll 1$, and so also provides a tiny contribution. Thus, the $|\uparrow\rangle$-projected Wigner function is dominated by the purely polaronic contribution, as we indeed demonstrate by the impressive agreement with the numerically exact NRG computation of $W^{(k)}_{\frac{1+\sigma^z}{2}}(X)$ in the left panel of Fig. 4.

In order to highlight the emergence of antipolaronic contributions in the wavefunction, we now insert the off-diagonal σ^+ Pauli matrix, which leads to the equivalent expression as defined in Eq. (7) of the main text:

$$W^{(k)}_{\sigma^+}(X) = \int \frac{\mathrm{d}^2\lambda}{\pi^2}\, e^{X(\bar{\lambda}-\lambda)} \langle GS|e^{\lambda a_k^\dagger - \bar{\lambda} a_k}\sigma^+|GS\rangle. \tag{32}$$

We thus need to compute

$$\begin{aligned} W^{(k)}_{\sigma^+}(X) = \frac{1}{2}\int \frac{\mathrm{d}^2\lambda}{\pi^2}\, e^{X(\bar{\lambda}-\lambda)} \Big[&\left\langle - f_q^{\text{pol.}}\middle|e^{\lambda a_k^\dagger - \bar{\lambda} a_k}\middle|f_q^{\text{pol.}}\right\rangle + p_2\left\langle - f_q^{\text{pol.}}\middle|e^{\lambda a_k^\dagger - \bar{\lambda} a_k}\middle|f_q^{\text{anti.}}\right\rangle + p_2\left\langle - f_q^{\text{anti.}}\middle|e^{\lambda a_k^\dagger - \bar{\lambda} a_k}\middle|f_q^{\text{pol.}}\right\rangle \\ &+ p_2^2\left\langle - f_q^{\text{anti.}}\middle|e^{\lambda a_k^\dagger - \bar{\lambda} a_k}\middle|f_q^{\text{anti.}}\right\rangle \Big]. \end{aligned} \tag{33}$$

A computation similar to performed above leads to the final result:

$$W^{(k)}_{\sigma^+}(X) = \frac{1}{\pi} e^{-2\sum_{q\neq k}(f_q^{\text{pol.}})^2} e^{-2X^2} + \frac{p_2}{\pi} e^{-\frac{1}{2}\sum_{q\neq k}(f_q^{\text{pol.}}+f_q^{\text{anti.}})^2} \left[e^{-2\left(X-\frac{f_k^{\text{pol.}}-f_k^{\text{anti.}}}{2}\right)^2} + e^{-2\left(X+\frac{f_k^{\text{pol.}}-f_k^{\text{anti.}}}{2}\right)^2} \right] + \frac{p_2^2}{\pi} e^{-2\sum_{q\neq k}(f_q^{\text{anti.}})^2} e^{-2X^2} \ . \tag{34}$$

The above expression shows important differences from the $|\uparrow\rangle$-projected Wigner distribution of Eq. (31). Indeed, the first term associated with the purely polaronic response is now of order $e^{-2\sum_{q\neq k}(f_q^{\text{pol.}})^2} \simeq \left\langle f_q^{\text{pol.}} \middle| -f_q^{\text{pol.}} \right\rangle \simeq \Delta_R/\Delta \ll 1$, and so is subdominant to the second contribution (with mixed polaron-antipolaron origin) of order $p_2 \propto \Delta/\omega_c$ (Note that the overlap $e^{-\frac{1}{2}\sum_{q\neq k}(f_q^{\text{pol.}}+f_q^{\text{anti.}})^2}$ appearing in the second term is of order 1). The third term, of order $p_2^2\Delta_R/\Delta$ is even more smaller. Thus, the off-diagonal Wigner function $W^{(k)}_{\sigma^+}(X)$ can be used to highlight the emergence of antipolarons in the many-body ground state wavefunction of the continuous spin-boson model, as was also discussed in the main text.

---

[1] Leggett, A. J., Chakravarty, S., Dorsey, A. T., Fisher, M. P. A., Garg, A. & Zwerger W. Dynamics of the dissipative two-state system. *Rev. Mod. Phys.* **59**, 1 (1987).

[2] Weiss, U. *Quantum Dissipative Systems* (World Scientific, 1993).

[3] Silbey, R. & Harris, R. Variational calculation of the dynamics of a two level system interacting with a bath. *J. Chem. Phys.* **80**, 2615 (1984).

[4] Chin, A. W., Prior, J., Huelga, S. F. & Plenio, M. B. Generalized polaron ansatz for the ground state of the sub-ohmic spin-boson model: An analytic theory of the localization transition. *Phys. Rev. Lett.* **107**, 160601 (2011).

[5] Nazir, A., McCutcheon, D. P. S. & Chin, A. W. Ground state and dynamics of the biased dissipative two-state system: Beyond variational polaron theory. *Phys. Rev. B* **85**, 224301 (2012).

[6] Braak, D. Integrability of the Rabi model. *Phys. Rev. Lett.* **107**, 100401 (2011).

[7] Raimond, J. M. & Haroche, S. *Understanding the Quantum* (Oxford Graduate Series, 2006).